REMARKS

Claims 1-26, 29-31 and 50-61 are currently before the Examiner. Claims 1-6, 9-13, 17, 19, 21, 23, 25, 26, 29, 30, 50 and 58 are amended herein. Claim 59 is cancelled without prejudice. New claims 62-64 are added. The amendments add no new matter.

Objection to the Claims:

Claims 10-18, 23-26 and 29-31 are objected to under 37 C.F.R. 1.75(c) as being in improper form because a multiple-dependent claim cannot depend from another multiple-dependent claim. Applicant submits that claims 10-13, 17, 23, 25, 26, 28 and 29 are amended herein to remove dependence from multiple dependent claim 8 (which is dependent from claims 1, 2 or 3), and that these amendments are sufficient to overcome this objection. New claim 62 is added to recapture material deleted from claim 29 by the removal of the multiple dependency. Similarly, new claims 63 and 64 are added to recapture material deleted from claim 30 by the removal of the multiple dependency. The amendments add no new matter.

Rejections under 37 C.F.R. §112, First Paragraph:

Claims 1-9, 19-22 and 50-61 are rejected under 35 U.S.C. §112, first paragraph for lack of enablement.

First, the Office Action states that "claims 1, 2 and 3 are directed to methods of using a computer system to identify a microbe inhabiting a host organism. The final step in each claim recites 'wherein absence …. indicates that said at least one sequence is a candidate sequence belonging to a microbe.'" The Office Action states that the method steps as recited do not provide the information required by the goal of the preamble, and that "the identification of a candidate sequence does not mean that the sequence is in fact from a microbe nor does it identify what microbe it might be." Similar statements are made with respect to the preamble and outcome of claims 50 and 58.

Applicant submits that the amendment of claims 1, 2, 3, 50 and 58 herein is sufficient to overcome this ground of rejection. Specifically, the preamble of each of these claims has been amended to recite "A method of using a computer system to determine the presence of a microbe inhabiting a host organism, comprising the steps of..." In addition, the body of each of these claims has been amended to recite "thereby indicating the presence of a microbe inhabiting said host organism." Applicant submits that the preamble and outcome of the claims as amended are in agreement, and respectfully requests the withdrawal of this enablement rejection.

Also with regard to enablement, the Office Action states "[t]he specification discloses identification of foreign gene sequences by transcript filtering against a variety of databases containing known and identified (at least as to source) sequences. The Office Action further states "[n]ote that the critical features set forth in the specification filtering multiple times against databases known for their completeness (or presumed completeness such as the human and mouse genome databases) or against databases with known characteristics (such as repeated sequences) are not present in the claims."

With regard to the suggested omission of a critical feature, Applicant submits that "[i]n determining whether an unclaimed feature is critical, the entire disclosure must be considered. Broad language in the disclosure (including the abstract) omitting an allegedly critical feature tends to rebut the argument of criticality." *In re Goffe*, 542 F.2d 564, 567 (C.C.P.A. 1976), citing *In re Anderson*, 471 F.2d 1237 (C.C.P.A. 1973). "Also, features that are merely preferred are not critical." *In re Goffe* at 567, citing In re Armbruster, 512 F.2d 676 (C.C.P.A. 1975).

Applicant submits that "filtering multiple times" is not a "critical feature" of the claimed invention. First, the specification states at page 11, lines 16-23:

> "The invention provides a computational subtraction method for detecting and identifying microbe sequences. The method comprises comparing the sequence information of a plurality of sequences obtained from one or more host organisms with sequences in a genomic database of host sequences to identify which of the plurality of sequences are not found (i.e., do not match other sequences) in the database. Sequences which are not found in the database are identified as

> candidate sequences which are likely to belong to a microbe. Preferably,
> sequence information from sets of sequences (two or more sequences, and
> preferably ten or more sequences) are compared against the entire genomic
> database at a time.

This description of the invention provided in the specification does not include any requirement for multiple filtering steps. Applicant submits that under *In re Goffe* and *In re Anderson*, this description would rebut the suggested criticality of the multiple filtering steps. That is, the specification does not teach that the filtering steps are critical features of the claimed methods. The use of additional filtering steps is a preferred embodiment, as evidenced by their inclusion in dependent claim 25. Thus, under *In re Goffe* and *In re Armbruster*, these features that are merely preferred are not critical.

Second, the independent claims result in the indication that the at least one sequence absent from the plurality of host sequences "is a *candidate* sequence belonging to a microbe, thereby indicating the presence of a microbe inhabiting the host organism." These independent claims (1-3, 50) as amended do not recite the identification of a microbe, not do they require proof that the absent sequence belongs to a microbe - the claims result in and require that "at least one sequence is identified as a *candidate* sequence belonging to a microbe," which is thus indicative of the presence of a microbe inhabiting the host organism. The fact that the recited steps identify a "candidate" microbe sequence, rather than requiring absolute identification of the sequence as microbial, or as coming from a specific microbe, means that the claim need not recite "multiple filtering steps." That is, the steps as recited result in the identification of at least one (meaning more than one are possible) candidate sequence. While additional filtering steps as taught in the specification can be used to narrow the number of candidate sequences and more specifically indicate the presence of a microbe inhabiting the host organism, the recited step of "searching a database of host organism genomic sequences to determine the presence or absence of said plurality of sequences in said database" results in the identification of at least one candidate sequence that is indicative of the presence of a microbe. Applicant submits that the identification of at least one candidate sequence that indicates the presence of a microbe in a host

organism thus does not require "filtering multiple times" – the candidate microbe sequences identified by searching in a search of the host organism can be further filtered if necessary to further narrow the candidates and indicate the presence of a microbe inhabiting the host organism.

Applicant notes, as mentioned above, that while "filtering multiple times" is not a required aspect of the invention, preferred embodiments can involve one or more filtering steps, as recited, for example, in claim 25, which recites the elimination of one or more of vector sequences, repetitive sequences, mitochondrial sequences, non-host species sequences, known host organism sequences, and combinations thereof from the genomic database comprising sequences from the host organism.

Further regarding enablement, the Office Action states:

> Note that the claims have no limitations as to how presence or absence of a sequence is determined. That is, the specification discloses particular algorithms and cut-offs to define presence and absence whereas the claims taken on their face would seem to require exact matching for presence or absence. As such, if the sequence obtained in (a) was longer than a sequence in (b), it would not be considered to be present. Likewise, a sequence reflecting a strain difference would not be considered to be present."

Applicant submits that in view of the definitions in the specification for the terms "present," "absent" and "a match", the presence or absence of a sequence according to the claims is not limited to exactly matching sequences. Specifically, the specification states "As used herein, a query sequence is 'present' in a database if the database contains a sequence which matches the query sequence and is 'absent' in a database if the database does not contain the matching sequence." (Specification page 8, lines 21-23) Regarding the term "a match," as it is used in the application, and particularly as it is used in the definitions of "present" and "absent," the specification states "As used herein, a 'match' between sequences refers to a level of sequence similarity equivalent to a BLAST score ranging from 40 (the equivalent of 20 consecutive identical nucleotides) to

2000 (the equivalent of 1000 consecutive identical nucleotides." Regardless of the exact BLAST, FASTA, etc. algorithm(s) and alignment parameters used to align sequences, the specification thus teaches that a "match" encompasses sequences comprising a minimum of 20 consecutive identical nucleotides. A failure to identify at least 20 identical consecutive nucleotides in a sequence means the sequence is not a "match" according to the invention. Thus, reading the claims in light of the specific definitions of the claim terms in the specification, a broader meaning is set forth for the terms "present" and "absent" than that argued in the Office Action. Using this teaching in the specification and ordinary skill in the art, Applicant submits that one of skill in the art would be enabled to perform the claimed invention without undue experimentation. Applicant respectfully requests the withdrawal of this enablement rejection.

Further regarding enablement, the Office Action states that the specification "does not provide databases containing expressed sequence tags (ESTs) or genomic sequences for all host organisms embraced by the claims." The Office Action further states that

> "While the specification indicates that many such databases are available, there is
> no evidence that these databases possess the type of sequence integrity (produced
> from host organisms not having a particular pathogenic condition, see for example
> claim 8) required by the claims. For example, with respect to claim 2, genome
> databases for host organisms with known symbiotic organisms would reasonably
> be expected to contain at least some of the symbiotic sequences. The Relman
> reference acknowledges that microbial and viral transcripts will be present in
> what are ostensibly human EST databases. The specification also acknowledges
> that microbe sequences can be present due to contamination in preparing libraries
> or sequences rather than actually being present in the host organism."

From this, the Office Action concludes that claims 1-9, 19-22 and 50-61 are not enabled. Applicant respectfully disagrees.

First, Applicant submits that Example 1 of the specification provides a working example of the claimed invention. Example 1 demonstrates that HPV sequences present in a library of HeLa nucleic acid sequences are identified as candidate microbe sequences using a method

encompassed by the claims. Applicant submits that the likely presence of microbial sequence contamination, e.g., from *E. coli* or other cloning hosts, in the human genome database did not interfere with the successful performance of the method in identifying candidate HeLa sequences that turned out to be from HPV. Further, inventor Dr. Matthew Meyerson describes additional instances of the successful use of embodiments of the claimed invention in the Declaration under Rule 1.132 filed herewith. Specifically, Dr. Meyerson describes the use of an embodiment of the claimed invention to identify viral (EBV) sequences present in a sample of human lung tissue taken from a patient with post-transplant lymphoproliferative disorder (PTLD). The Declaration states that computational subtraction identified 10 sequences that "could be amplified only from Epstein-Barr virus (EBV)-infected tissues," and that "all 10 corresponded to known EBV sequences." The Declaration states that this experiment demonstrates that computational subtraction can detect pathogenic microbes in primary diseased tissue from humans despite the possible presence of microbial sequences in the human sequence databases."

Dr. Meyerson further states in his Declaration that he has not "performed any experiments on diseases of known microbial origin that have failed to detect the causative microbe." He further states:

> "the issue of microbial contamination of genomic sequence databases will diminish with the curation and re-sequencing of these databases. At present, the human genomic DNA databases do not contain detectable sequences from pathogenic microbes. Furthermore, the sequences of other microbes can also not be detected in the human genome sequence, to our knowledge. Other genomic sequence databases will rapidly approach the quality of the human genome sequence."

Dr. Meyerson's Declaration also states:

> "the libraries that we have generated from diseased human tissue have been completely free of microbial sequence contamination, thanks to careful experimental technique. At this point, we have generated over 100,000 unique sequence reads from libraries of human diseased tissue, not known to be caused by pathogenic infection. We have detected no microbial sequences in these libraries.

In view of Dr. Meyerson's additional experimental data and his statements regarding the minimal occurrence and impact, if any, of microbial sequence contamination of genomic databases, Applicants submit that claims 1-9, 19-22 and 50-61 are fully enabled by the specification.

Second, while it is possible that the presence of contaminating sequences from a microbe in, for example, a genomic database, may hinder the indication of the presence of that specific microbe in a host organism, this does not render the claimed method incapable of determining the presence of any, or even most or a great number of microbes that may inhabit a host organism. That is, the presence of contaminating microbial sequences in a host genomic sequence database may lead to some inoperable embodiments, but such embodiments are very few relative to the number of other microbes whose sequences do not contaminate the database(s). Further, the claims do not require the determination of the presence of *all* microbes in the host organism, but rather the indication of the presence of "*a*" microbe. While the presence of contaminating sequences may lead to a low background of false-negative results for certain kinds of microbes, e.g., cloning strains or vectors, Applicant submits that sequences for many other microbes are not likely present in genomic databases and will be identified as candidate microbial sequences according to the claimed methods.

The standard for enablement when there are possible inoperative embodiments encompassed by the claims is whether one of skill in the art would be able to determine an operative from an inoperative embodiment without undue experimentation. *Atlas Powder Co. v. E.I. DuPont de Nemours & Co.*, 750 F.2d 1569, 1577 (Fed. Cir. 1984). Applicant submits that in this instance such a determination is very straightforward, in that the identification of a sequence in the test database that is absent in the genomic sequence database indicates an operative embodiment. In view of the above, Applicant submits that the possible presence of some contaminating microbial sequences in genomic sequence databases useful in the claimed methods does not render the claimed invention non-enabled. Applicant respectfully requests the withdrawal of this enablement rejection.

Finally, with regard to enablement, the Office Action states "claims 1, 3 and 50 recite 'obtaining sequence information ... from at least one host organism." The Office Action asks "If sequence information is obtained from both a human organism and a *Drosophila* organism, none of the steps recited results in discrimination or assignment of any absent sequence from one or the other host organism or identification of a microbe inhabiting either one."

Applicant submits that the amendment of claims 1, 3 and 50 to recite "obtaining sequence information from a plurality of sequences from *a* host organism" is sufficient to overcome this rejection. Applicant notes that the language "*a* host organism" includes is inclusive of sequence information from one host organism, but the claim as amended also encompasses the situation in which the sequence information is obtained from more than one host organism of a given species. In view of this amendment, Applicant respectfully requests the withdrawal of this rejection of claims 1, 3 and 50.

Rejections under 35 U.S.C. §112, Second Paragraph:

Claims 4-6, 19-22, 51 and 52 are rejected as indefinite under 35 U.S.C. §112, second paragraph.

Claims 4-6 are said by the Office Action to recite limitations with respect to microbes. The Office Action states that "the method steps of the claims provide no steps directed to identifying these particular organisms," concluding that the claims are either not further limiting or confusing.

Applicant submits that the amendment of claims 4-6 herein is sufficient to overcome this rejection. Specifically, claim 4 is amended to recite "wherein said candidate sequence belongs to a symbiotic microbial organism, claim 5 is amended to recite "wherein said candidate sequence belongs to a mutualistic organism, a commensal organism, or a parasitic organism, and claim 6 is amended to recite "wherein said candidate sequence belongs to a pathogenic organism."

Claim 19 is said to be confusing in reciting limitations with respect to an "expressed sequence." The claim depends from claims 1, 2 or 3 through claim 8, but the Office Action states that claim 8 does not recite expressed sequences and neither do claims 1 or 2, and that it is not clear whether the limitation of claim 19 is directed to part (a) or (b) or both.

Applicant submits that the amendment to claim 19 made herein is sufficient to obviate this rejection. The dependency has been changed such that the claim is now dependent upon claims 3 and 10, which both recite "expressed sequences." Claim 19 as amended recites "[t]he method according to claim 3 or 10, wherein said expressed sequences are identified using a differential gene expression assay.

Claims 21 is said to be confusing in reciting "said at least one sequence." The claim depends from claim 8, which depends from claims 1, 2 or 3. The Office Action states that claims 1, 2 and 3 recite "at least one sequence" with respect to computational means (database comparison) and not experimental means (subtractive hybridization).

Applicant submits that the amendment to claim 21 made herein is sufficient to obviate this rejection. The dependency has been changed such that the claim is now dependent upon claims 3 and 10, and "at least one sequence" has been amended to recite "candidate sequence."

Claim 52 is said to appear to contain all of the limitations of claim 51 "as claim 50 already requires that the sequences are expressed sequences." The Office Action states that, "alternatively, claim 51 provides no steps to distinguish expressed from non-expressed sequences in the library."

Applicant submits that the cancellation herein of claim 52 is sufficient to obviate this rejection.

Applicant has also amended claim 9 to correct a lack of antecedent basis not mentioned in the Office Action. The term "said identified sequences" has been amended to "said plurality of sequences." The amendment adds no new matter.

In view of the above, Applicant respectfully requests the withdrawal of the §112, second paragraph rejection of claims 4-6, 19-22, 51 and 52.

Rejections under 35 U.S.C. §102(b):

Claims 1, 3-6, 50-53, 58 and 59 are rejected under 35 U.S.C. §1-2(b) as being anticipated by Sabatini et al. (U.S. Patent 5,966,712). The Office Action states that Sabatini et al. discloses making nucleic acid libraries from various host organisms, obtaining sequence information therefrom (including expressed sequences), and searching other databases to determine unique or common sequences." The Office Action continues "computational subtractions and Southerns are disclosed" and "microbial libraries are specifically disclosed. The Office Action further states "note that although the preamble of the claims is directed to identifying a microbe, the method steps only require obtaining sequence information and searching a database to determine presence or absence of sequences." The Office Action thus concludes that the reference anticipates the claims. Applicant respectfully disagrees.

Applicant submits that Sabatini et al. does not teach obtaining sequence information from a plurality of sequences from at least one host organism and searching a database of host organism genomic sequences to determine the presence or absence of the plurality of sequences in the database, wherein the absence of at least one of the sequences in the database indicates that the at least one sequence is a candidate sequence belonging to a microbe, thereby indicating the presence of a microbe inhabiting the host organism. Specifically, Applicant submits that there is no teaching of the indication of the presence of a microbe inhabiting a host organism in the Sabatini et al. reference. The Sabatini et al. reference provides no teaching at all regarding the use of a computer system to determine the presence of a microbe inhabiting a host organism. In view of this, and because each of independent claims 1, 3, 50 and 58 recites a method for such determination, Applicant submits that Sabatini et al. cannot anticipate any of claims 1, 3, 50, 58 or claims 4-6, 51-53 and 59 that depend from them.

Claim 58 is rejected under 35 U.S.C. §102(b) over Fredricks & Relman, 1999, CID 29: 475-488. The Office Action states that "Fredricks et al. took a sample of CSF to provide DNA from a human patient" and that "[t]he DNA was amplified using PCR and the amplified sequences compared to a database of known microbial 16S rRNA to identify the microbe." The Office Action thus concludes that the reference anticipates claim 58. Applicant respectfully disagrees.

Applicant submits that the Fredricks & Relman reference does not teach "searching a database of microbial sequences from a library of expressed sequences to determine the presence or absence of said plurality of sequences from at least one non-microbial host organism in the database, wherein the presence of an expressed sequence from said at least one non-microbial host organism in the database identifies the expressed sequence as a candidate microbe sequence, thereby indicating the presence of a microbe inhabiting said host organism" as required by claim 58 as amended. Specifically, the Fredricks & Relman reference does not teach searching a database of microbial sequences from a library of expressed sequences. Support for the amendment is found in the specification and in original claim 59, which is cancelled herein. As such, Applicant submits that the Fredricks & Relman reference does not anticipate claim 58 as amended. Applicant respectfully requests withdrawal of the §102(b) rejection of claim 58 over this reference.

In view of the above, Applicant respectfully requests the withdrawal of the §102 rejections of claims 1, 3-6, 50-53, 58 and 59 over Sabatini et al., and of claim 58 over Fredricks and Relman.
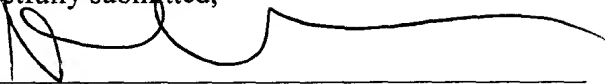
In view of the above, Applicant submits that all issues raised in the Office Action have

been addressed herein. Applicant respectfully requests reconsideration of the claims.

Respectfully submitted,

Date: _____7\10\03_____

_____
Kathleen M. Williams, Ph.D.
Registration No.: 34,380
Palmer & Dodge LLP
111 Huntington Avenue
Boston, MA 02199-7613
Telephone: (617) 239-0100
Telecopier: (617) 227-4420